View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by UMP Institutional Repository

Predefined Object Reduction

1 **Mohammed Adam TaheirMohammed,1Wan Maseri Binti Wan Mohd, 1Ruzaini Bin Abdullah Arshah, 2LiuYao**

***Abstract* - Reduction techniques is still an open area to be explored in knowledge management. This paper defines algorithm known as Predefined Hybrid Reduction which generate its conditions for object co occurrences of original data then execute Hybrid Reduction data for their data to perform extractions. Predefined Hybrid Reduction give a proper solution for expansion the data set , it select significant object with high quality of informations, it delete every object not satisfies their conditions. It show appropriate relevant result. It provide better reduction without inconsistency problem unlike data comparisons. It manage the inferior object which store only significant data based on predefined confidence and predefined support for maintain the inferior object then Hybrid reduction which are dual reduction. As part of this proposal, a comparison test with Hybrid reduction. The conclusion part which shows better alternative result through our model.**

***Keywords*: Boolean-valued information system; Extractions reductions; Parameters reductions ;Knowledge Management.**

I. Introduction

Handling uncertain data solved by using mathematical principles, and one of them is soft set theory [2]. Soft sets are called (binary, basic, elementary) neighborhood systems. As for standard soft set," it may be redefined as the classification of objects in two distinct classes, thus confirming that soft set can deal with a Boolean-valued information system". Molodtsov [2] advantages it is free from the inadequacy of the parameterization tools, unlike in the theories of fuzzy set, probability and interval mathematics. The knowledge management requires effective knowledge organization, searching and sharing strategies. The problems are how to analyze large amount of data[7]and if its domain is not able to extracting their data.

In recent years, research on soft set theory has been active using the fundamental soft set theory, soft set theory in abstract algebra and soft set theory for data analysis in decision making [2,9,10]. The concept of soft set-based reduction is another area which purportedly supports decision making problems with less involvement of data and attributes by reducing the attributes. The objective of this paper to find better reduction and at the same time managing uncertain data by remove the objects ambiguity which provide high quality for information retrieval which avoiding similar classes(duplicated) or no needed information (vague) with proper classification and organized in efficient manner. Querying data set and expansion their relevant retrieval managing the query storage by prober reductions which are high degree of data independence based on soft set theory. To perform proper expansion and at the same time avoiding the inferior object problems if its high , this techniques saving the researcher search time and satisfies their guessing.

The rest of this paper is organized as follows. Section III describes the fundamental concept of soft set theory. Section IV presents analysis of Hybrid reduction. Section V is a proposal techniques which is based on Hybrid reduction techniques reductions followed by section VI which focuses on result and discussions. Finally, the conclusion of this paper is described in section VII.

II. Related work

Maji et al. [1] in the year 2002 introduces techniques for extract the data by generating optimal and sub optimal decision, but its sup optimal result are not correct [8]. To perform a reduction one way is to remove uncertain data by prober relation which manage the duplicated data as well as find the relation between object which are significant un like the data comparisons. The soft set function mapping their concepts for reduce the large amount of data with proper reductions and high dependency between objects. Chen et al. techniques [9] solve the problems in [1] in the year 2005 by removing the inconsistency from it, but not mentioned the sub optimal extractions [8]. Kong et al. [10] in the year 2008 solve [9] problem by normal parameter reduction, but has implies problems are hard to applies and not work on our data set table 1 [8]. Rose et al. [11] in the year 2010 defines techniques which overcome Maji [1] problem, it investigate that every sub parameters combination which satisfies the optimal result partions and has support cluster same as the original data set support cluster, but its

All Rights Reserved © 2013 IJARCET

problem has low reduction [12]. All these techniques mentioned above focuses on column reduction to this Rose et al. [5] in the year 2011 applies reduction techniques for row reduction as well as column reductions which select the object that not in the maximum weights and has parameters are extracted same as [11]. This proposal based on Hybrid reductions, it applies predefine confidence and predefined support to Hybrid reductions for managing the inferior object to remove vague objects from extraction to this data size reduce as well as the significant object are retrieves by has prober association and reduction.

III. Soft Set Theory

Throughout this section U refers to an initial universe, E is a set of parameters, $P(U)$ is the power set of U.

Definition . (See [2].) A pair (F,E) is called a soft set over U, where F is a mapping given by

$$F : E \to P(U) \qquad (1)$$

Throughout this section U refers to an initial universe, E is a set of parameters, $P(U)$ is the power set of U.

Definition .. (See [2].) A pair (F,E) is called a soft set over U, where F is a mapping given by

$$F : E \to P(U) \qquad (2)$$

In other words, a soft set is a parameterized family of subsets of the universe U. For $\varepsilon \in E$, $F(\varepsilon)$ may be considered as the set of ε-elements of the soft set (F,E) or as the set of ε-approximate elements of the soft set, instead of a (crisp) set.

Example as mentioned in [13] Let a soft set (F,E) representing studies the communication prowess among selected university student. Let assume that there are eighteen students that has been surveyed in the universe U with $U = \{u_1, u_2, \cdots, u_{18}\}$, and E is a set of parameters representing communication facilities that is been used by the student surveyed, $E = \{p_1, p_2, p_3, p_4, p_5, p_6\}$, where p_1 stands for the parameter for using communication facilities such as "email", p_2 stands for the parameter "facebook", p_3 stands for the parameter "blog", p_4 stands for the parameter "friendsters", p_5 stands for the parameter "yahoo messenger" and lastly p_6 stands for the parameter "sms" .

Consider the mapping $F : E \to P(U)$ given by "student communication prowess $(\cdot)$", where $(\cdot)$ is to be filled in by one of parameters $p \in E$.

Suppose that

As for example, $F(p_2)$ means communication by facebook is been used by and being represented functional value of $\{u_2, u_3, u_4, u_5, u_8, u_9, u_{10}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}\}$, while $F(p_4)$ means communication through friendsters with its functional value represented by $\{u_2, u_3, u_4, u_8, u_9, u_{10}, u_{11}, u_{13}, u_{15}, u_{16}, u_{17}, u_{18}\}$.

Thus, the overall approximation can be represented as the following:

$$(F,E) = \left\{ \begin{array}{l} \text{email} = \{u_1, u_2, u_3, u_4, u_8, u_9, u_{10}, u_{11}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}\}, \\ \text{facebook} = \{u_2, u_3, u_4, u_5, u_8, u_9, u_{10}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}\}, \\ \text{blog} = \{u_1, u_2, u_3, u_4, u_6, u_8, u_9, u_{10}, u_{11}, u_{13}, u_{15}, u_{16}, u_{17}, u_{18}\}, \\ \text{friendsters} = \{u_2, u_3, u_4, u_8, u_9, u_{10}, u_{12}, u_{13}, u_{15}, u_{16}, u_{17}, u_{18}\}, \\ \text{ym} = \{u_2, u_3, u_4, u_8, u_9, u_{10}, u_{11}, u_{13}, u_{15}, u_{16}, u_{17}, u_{18}\}, \\ \text{sms} = \{u_2, u_3, u_4, u_8, u_9, u_{10}, u_{13}, u_{15}, u_{16}, u_{17}\} \end{array} \right\}$$

Figure 1. The soft set

The previous example can be represented in the following Boolean-valued information system

Table 1. Tabular representation of a soft set

U/P	p_1	p_2	p_3	p_4	p_5	p_6	$f(\cdot)$
u_1	1	0	1	0	0	0	2
u_2	1	1	1	1	1	1	6
u_3	1	1	1	1	1	1	6
u_4	1	1	1	1	1	1	6
u_5	0	1	0	0	1	0	2
u_6	0	0	1	0	1	0	2
u_7	0	0	0	0	1	0	1
u_8	1	1	1	1	1	1	6
u_9	1	1	1	1	1	1	6
u_{10}	1	1	1	1	1	1	6
u_{11}	1	0	1	0	1	0	3
u_{12}	0	1	0	1	0	0	2

All Rights Reserved © 2013 IJARCET

u_{13}	1	1	1	1	1	1	6
u_{14}	1	1	0	0	0	0	2
u_{15}	1	1	1	1	1	1	6
u_{16}	1	1	1	1	1	1	6
u_{17}	1	1	1	1	1	1	6
u_{18}	0	0	1	1	1	0	3

Table 1 continuation

IV. Analysis of Hybrid Reduction in Soft Set Decision Making in Rose et al [5]

A. The idea of Rose et al. [5] :

1- Input soft set (F,E) over universe U.
2- Determine co-occurrences parameter in each object.
2- Calculate the support for every object .
3-Determine the order of supports by arranging in decreasing order.
4- Rank the object based on optimal object ,sub optimal ,next obtimal until the inferior object which determine based on support.
5- Determine the U/E clusters partions.
6- Determine any group of attributes satesfies optimal decision partions and suppA(u)=supp(v) for every u,v in U.
7- For any group of attributes determine the reduction form .
8-Determine any row fulfill the definition of ultimate support set.
8- Delete ultimate support partions.
9-For any ultimate minimum support delete the partions of inferior object.
10-If there is any ultimate minimum support set , mark the mark the object the inferior object.
11- Remove every row or columns which has empty objects(zero significant).

The parameters co-occurrence set is the representation of the value 1 which as Coo u1 ={p1 ,p3} until last object.

The parameters co-occurrence set is the representation based on the value 1 which introduce co occurrences like Coo u1 ={p1 ,p3} then the weight for every supp based on table 1 as follow and the result of Hybrid reductions shown as in table 2.

Supp (u i) =6 i u , i =2,3,4,8,9,10,13,15,16,17
Supp (u j) =3 j u , j =11,18
Supp (u k) =2 k u , k =1,5,6,12,14
Supp(u L) = l , L =7

Table 2

U/P	p_1	p_2	p_3	p_4	p_5	p_6	$f(.)$
u_1	1	0	1	0	0	0	2
u_5	0	1	0	0	1	0	2
u_6	0	0	1	0	1	0	2
u_7	0	0	0	0	1	0	1
u_{11}	1	0	1	0	1	0	3
u_{12}	0	1	0	1	0	0	2
u_{14}	1	1	0	0	0	0	2
u_{18}	0	0	1	1	1	0	3

V. The Proposal techniques

In this section, an alternative concept of object reduction based on predefined support and predefined confidence are introduced. The main idea behind the object reduction is to further reduce the size of database without compromising on the values of objects sub-optimal decisions, or even the next sub-optimal decision of objects. For a Predefined Hybrid Reduction , it is proposed on object reduction which remove false frequent object occurrences that not satisfies Predefined conditions. To this, the notion of Predefined Hybrid reduction is presented firstly then Hybrid Reduction are executed which increase response time. This techniques has dual reduction, first the object data size are reduce by Predefind Hybrid Reduction, second the inferior object and inconsistency are removed by Hybrid Reduction. This techniques based on rows reductions and it proposed to maintain the object reduction.

A. proposal procedure as follows:

. The procedure of Predefined Hybrid reduction are:

1- Take the input from Hybrid reduction.
2- Calculate every object co occurrences (support).
3- Determine optimal result partions
4- Calculate the confidence co occurrences for every object not in optimal result partions.

All Rights Reserved © 2013 IJARCET

5- If object support < (predefined support and predefined confidence) then the object are deleted .
6- Execute Hybrid reduction techniques.
7- This procedure known as Predefined Hybrid Reduction.
8- Apply Predefined Hybrid Reduction for every sub parameters generated by Hybrid reduction.

B .Analysis of Predefined Hybrid reduction

The association rule which safeties the certain constrain are min support and min confidence mentioned in [13].

In table 3 Coo(u1)=p1,p3, Coo(u5)=p2,p5, Coo(u6)=p3,p5 , Coo(u7)=p5, Coo(u11)=p1,p3,p5, Coo(u12)=p2,p4
Coo(u14)=p1,p2, Coo(u18)=p3,p4,p5

Now with help of parameter co-occurrences we calculated the support as follow:

Sup(u1)= p1,p3 thus conf p1→ p3=2/3 = 67%
Sup(u5)= p2,p5 thus conf p2 → p5=1/3 = 33%
Sup(u6)= p3,p5 thus conf p3 → p5=3/4 = 75 %
Sup(u7)= p5 thus conf p5 =0.0 %
Sup(u11)= p1,p3,p5 thus conf p1,p3→ p5 1/2 = 50%
Sup(u12)= p2,p4 thus conf p2→ p4 =1/3= 33%
Sup(u14)= p1,p2 thus conf p1→ p2=1/3=33%
Sup(u18)=p3,p4,p5 thus conf → p3,p4 p5=1/1=100%

Table 3

U/P	p_1	p_2	p_3	p_4	p_5	$f(.)$
u_1	1	0	1	0	0	2
u_5	0	1	0	0	1	2
u_6	0	0	1	0	1	2
u_7	0	0	0	0	1	1
u_{11}	1	0	1	0	1	3
u_{12}	0	1	0	1	0	2
u_{14}	1	1	0	0	0	2
u_{18}	0	0	1	1	1	3
supp	3	3	4	2	5	

Table 3 continuation

suppose min supp is 2, and min confident is 35%. Therefore the object u5 are not satisfies the predefined confidence 35% and predefined support thus the object u5 is deleted from data set, the result of Predefined Hybrid Reduction as shown in table4.

Table 4

U/P	p_1	p_2	p_3	p_4	p_5	$f(.)$
u_1	1	0	1	0	0	2
u_5	0	1	0	0	1	2
u_6	0	0	1	0	1	2
u_{11}	1	0	1	0	1	3
u_{12}	0	1	0	1	0	2
u_{14}	1	1	0	0	0	2
u_{18}	0	0	1	1	1	3

VI.Result and discussions

The Predefiend Hybrid Reduction are requeried predefind confidence and predefined support. It generates its role association for object reduction as mentioned in [13] the response time are increased compared to Hybrid Reduction which are eliminated (deleted) from the data set, moreover the frequent object co occurances affects the information perecision and recall raio. Every object not satisfies predefind support and predefind confidences are deleted from data set directly which increase the objects reduction size.

The Hybrid reduction store the proper reduction of table which occupies 73% of memmory size, while Predefiend Hybrid Reduction based on that predefined confidence reduce Hybird reduction data size which offer more free memory size compared to Hybrid reduction. It the object which not satifies the Predefined Hybrid reduction conditions.

VII..Conclusion

The Predefiend Hybrid Reduction proposed for object reduction, it requierd preefind support and predefind confidence for object extraction which are satesfies their conditions. By using Predefiend Hybrid Reduction frequent object extraction are enhanced the data set size reduction to this uncertain data (vague information) are manged. Predefiend Hybrid Reduction show better result

All Rights Reserved © 2013 IJARCET

compared to Hibrid Reduction for maintaining the inferior object which increased the responce time. This approach reduced the number of objects in Boolean databases drastically but still been able to maintain consistency in decision making.

References

[1] P.K.Maji, A.R.Roy, and R.Biswas 2002. An application of soft sets in a decision making problem, Computer and Mathematics with Application 44, 1077–1083.

[2] D.Molodtsov 1999. Soft set theory-first results, Computers and Mathematics with Applications 37, 19–31

[3] Y.Zhao, F.Luo, S.K.M.Wong, and Y.Y.Yao 2007. A general definition of an attribute reduct, Lecture Notes in Artificial Intelligent 4481, 101–108.

[4] Z.Pawlak, and A.Skowron 2007. Rudiments of rough sets, Information Sci2ences 177 (1), 3–27 .

[5] A.N.M.Rose, M.I.Awang, H.Hassan, A. H.Zakaria, T.Herawan, and M.M.Deris 2011. Hybrid Reduction in Soft Set Decision Making , pp. 108–115, 2011.© Springer-Verlag Berlin Heidelberg 2011.

[6] J.Groppe , S.Groppe, A.Schleifer 2011. Visual Query System for Analyzing Social Semantic Web , March 28–April 1, 2011, Hyderabad, India.

[7] P.A.K Jeffrey. F.HudekIhab. I.G.Weddell 2012. Interpreting Keyword Queries over Web Knowledge Bases , CIKM'12, October 29–November 2, 2012, Maui, HI, USA.Copyright 2012 ACM 978-1-4503-1156-4/12/10 .

[8]M.A.T.Mohammed,W.M.B.W.Mohd, R.B.A.Arshah, L.Yao 2013. Parameter Reduction Comparisions (under review).

[9] D.Chen, E.C.C.Tsang, D.S.Yeung, and X. Wang 2055. Computers and Mathematics with Applications .49(2005) 757–763 .

[1O] Z.Kong, L.Gao, L.Wang, and S.Li 2008. Computers and Mathematics with Applications .56(2008) 3029–3037

[11] A.N.M.Rose, T.Herawan and M.M.Deris 2010. A framework of Decision Making Based on Maximal Supported Sets. ISNN 2010, Part I, LNCS 6063, pp. 473–482, 2010. © Springer-Verlag Berlin Heidelberg 2010.

[12]M.A.T.Mohammed, W.M.B.W.Mohd, R.B.A.Arshah, L.Yao 2013. Comparison of Hybrid Reduction Technique with several parameter reduction techniques (under review).

[13]V.Rajpoot, Prof. S.k.Shrivastava, Prof. A.Mathur 2012. An Efficient Constraint Based Soft Set Approach for Association Rule Mining. IJERA, ISSN: 2248-9622, Vol. 2, Issue4, July-August 2012, pp.2210-2215.

All Rights Reserved © 2013 IJARCET